DOCUMENT RESUME

ED 253 556 TM 850 095

AUTHOR Marsh, Herbert W.
TITLE Global Self-Esteem: Its Relation to Weighted Averages of Specific Facets of Self-Concept.
PUB DATE 15 Jan 85
NOTE 29p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Correlation; *Factor Structure; Foreign Countries; *Predictor Variables; *Self Concept Measures; *Self Esteem; Weighted Scores
IDENTIFIERS Australia; *Self Description Questionnaire III

ABSTRACT

Theory and common sense posit that the effect of a specific facet of self-concept on Global Esteem will vary with the importance placed on that facet, but no support for this interactive hypothesis was found. Unweighted averages of 12 distinct dimensions of self-concept from the Self Description Questionnaire III correlated about .7 with Global Esteem, but weighting each facet by the importance assigned to it by the entire group, by diverse subgroups, or by each individual made no difference. Even "random number" weights did almost as well, while optimal weights derived from multiple regression did only slightly better, suggesting that differential weighting has little effect. Nevertheless, weak support for the hypothesis was found for the Spiritual and Physical Abilities facets, and these were the two facets where the importance ratings varied the most. Though Global Esteem was reasonably well predicted by the specific facets, few specific facets were adequately predicted by Global Esteem, arguing against the sole reliance on a single global measure in self-concept research. (Author)

Global Self Esteem: Its Relation to Weighted Averages of Specific Facets of Self-concept

Herbert W. Marsh

Department of Education

The University of Sydney

15 January, 1985

Running Head: Self-concept

Reprint requests should be sent to Dr. Herbert W. Marsh, Department of Education, University of Sydney, NSW 2006, Australia.

Global Self Esteem: Its Relation to Weighted Averages of Specific Facets of Self-concept

ABSTRACT

Theory and common sense posit that the effect of a specific facet of self-concept on Global Esteem will vary with the importance placed on that facet, but no support for this interactive hypothesis was found. Unweighted averages of 12 distinct dimensions of self-concept from the Self Description Questionnaire III correlated about .7 with Global Esteem, but weighting each facet by the importance assigned to it by the entire group, by diverse subgroups, or by each individual made no difference. Even "random number" weights did almost as well, while optimal weights derived from multiple regression did only slightly better, suggesting that differential weighting has little affect. Nevertheless, weak support for the hypothesis was found for the Spiritual and Physical Abilities facets, and these were the two facets where the importance ratings varied the most. Though Global Esteem was reasonably well predicted by the specific facets, few specific facets were adequately predicted by Global Esteem, arguing against the sole reliance on a single global measure in self-concept research.

Global Self Esteem: Its Relation to Weighted Averages of Specific Facets of Self-concept

The Multidimensionality of Self-concept.

Self-concept is widely posited to be a desirable outcome and to explain overt behaviors and other constructs in many areas of psychology. Despite the theoretical and practical significance of the self-concept construct, reviews of self-concept research typical identify a lack of theoretical models for defining and interpreting the construct, and the poor quality of measurement instruments used to assess it (Burns, 1979; Welles & Marwell, 1976; Wylie, 1974; 1979). In an attempt to remedy this situation, Shavelson, Hubner and Stanton (1976) reviewed existing theoretical and empirical research, and on the basis of this review posited a multifaceted hierarchical model of self-concept. Shavelson specified particular facets of self-concept and a hierarchical ordering among these facets; general self-concept, the apex of the hierarchy, was broken into academic and nonacademic components and each of these were subdivided into more specific facets (e.g., specific academic content areas, social, physical and emotional).

Initially Shavelson's multifaceted model was not widely accepted as many researchers viewed self-concept as either a unidimensional construct or one so heavily dominated by a general factor that separate components could not be readily differentiated. Through the mid-1970's self-concept instruments typically consisted of a hodge-podge of self-referent items, little effort was made to develop/refine these instruments in order to measure specific facets, and support for the multidimensionality of self-concept was weak. More recently, researchers have developed self-concept instruments specifically to measure particular facets of self-concept that are at least loosely based on an explicit theoretical model, and then used factor analysis to test the existence of these a priori facets. This approach has produced instruments in which multiple facets of self-concept are clearly identified and quite distinct (e.g., Boersma & Chapman, 1979; Dusek & Flaherty, 1981; Fleming & Courtney, 1984; Harter, 1982; Marsh, Smith, Barnes & Butler, 1983; Marsh, Barnes, Cairns & Tidman, 1984; Marsh, Barnes & Hocevar, in press; Soares & Soares, 1982).

Shavelson and Marsh (in press; Marsh & Shavelson, 1984) reviewed research stimulated by Shavelson's model and the Self Description Questionnaire (SDQ) instruments that were derived from it. They found strong support for the multidimensionality of self-concept, and the facets posited by the model. However, the hierarchical ordering was more complicated, and weaker, than originally suspected (also see Marsh, 1984); for late

adolescents the different facets of self-concept were relatively uncorrelated. They concluded that self-concept cannot be adequately understood if its multidimensionality is ignored.

General Self and Specific Facets of Self.

Historically self-concept research has emphasized a general, overall or total self-concept, and specific facets of the construct have been relegated a minor role. With the increased emphasis on the multidimensionality of self-concept, the specific facets have become more important, and the role of general self-concept has become less clear. There is no widely accepted definition of how the general construct should be defined and at least four operational definitions are common: a) a hierarchical general self that appears at the apex of hierarchical models such as Shavelson's model; b) a conglomerate general self that is the total score from a large hodge-podge of self-referent items that attempt to sample broadly from a range of self-relevant characteristics; c) a global self esteem scale that is relatively unidimensional and content-free in that it is comprised of items that infer a general sense of self-worth or self-confidence that could be applied to many specific areas (e.g., the General Self scale from the SDQ III, and other scales described by Rosenberg, 1965; 1979; and Harter, 1982); and d) a weighted average general self where specific facets are weighted according to their salience, value or importance (e.g., Hodge & McCarthy, 1984; Watkins, 1978). Three of these, all but global esteem, are derived from some *a priori* or empirically weighted average of specific facets, though neither the facets and nor their weights are not adequately specified in the conglomerate general self.

These alternative conceptualizations of the general self construct have implications for the role of the importance, salience or centrality of a specific area of self-concept in determining general self-concept. Historically, William James (1890/1963) argued that failure in areas deemed to be unimportant to a particular individual have little impact on general self, and this contention has been reiterated by many theorists (e.g., Coopersmith, 1967; Harter, 1982; 1983; Hodge & McCarthy, 1984; Maxwell & Welles, 1976; Rosenberg, 1965; 1979; Watkins, 1978; Wylie, 1974). Coopersmith (1967) indicated that an individual's self-appraisals might vary in different areas so that "his overall appraisal of his abilities would presumably weight these areas according to their subjective importance enabling him to arrive at a general level of self-esteem (p, 6). He went on to state that: "Though this appears to be the case, objective evidence on the method of arriving at general appraisals is sparse" (p. 6). Wylie (1974, p. 48) stated that: "The sum is simple expedient in the face of ignorance

and should be so recognized. Steps should be taken to weight item ratings according to their perceived salience to S, but this has not yet been tried". Rosenberg (1965; 1979; also see Hodge & McCarthy, 1984) proposed an interactive hypothesis whereby a positive self-perspective in a specific area seen as important contributes postively to Global Esteem, a negative self-perspective in an important area detracts from Global Esteem, and self-perspectives in unimportant areas have little impact on Global Esteem.

Theorists have also speculated about how individuals determine the importance of specific areas of self-concept, and how this relates to Global Esteem. Following Festinger's social comparison theory, Wells and Marwell (1976, p. 55) describe the dilemma of a person with a poor ability in a particular area who must balance a reality principal against the need to have a favorable self-evaluation. One possible compromise is to recognize the poor ability in the particular area, but to give this little importance in the determination of Global Esteem. Rosenberg (1982; p. 538) found support for a selectivity hypothesis in that an individual "will be disposed to value those things at which he considers himself to be good and to devalue those qualities at at which he considers himself poor" but he also recognized that "the freedom to select one's values in a fashion congenial to one's self-image is not, of course, without limit." Hodge and McCarthy (1984), and others, emphasize the constraints that group and subgroup values place on the freedom to select one's values. These may force individuals to place a high value on some facet where their self-perceptions are poor, thus leading to a lower Global Esteem.

The Hodge and McCarthy 1984 Study.

Hodge and McCarthy (1984) examined relations among specific facets of self-concept, their perceived importance, and Global Esteem (the Rosenberg scale) in a large sample of high students. Using a variety of empirical and a priori weightings of the specific facets, they found that no weighted average correlated with Global Esteem higher than .45, suggesting to them that the weighted summations measure a different construct than Global Esteem. An unweighted average of the specific dimensions correlated about .4 with Global Esteem, and there was little or no improvement when the specific facets were weighted by the mean importance rating assigned by the entire group, by students within the same school, or by students within the same class. Paradoxically, when each component was weighted by the individual's own rating of its importance, the weighted average was significantly less correlated (r=.34) with esteem than any other weighted or unweighted average. In other analyses they found little or no support for Rosenberg's interactive hypothesis. These findings led the authors to conclude, at least for this age group, that the group values are much more

important than individual values in determining the value placed on specific facets, and this was their most important theoretical conclusion. However, a number of methodological and theoretical issues require further consideration.

The Use of Single Item Rating Scales. Self-esteem was based on a multi-item scale, but the specific facets and their importance were inferred from single-item responses. Hence, while esteem had a marginally acceptable reliability (alpha = .75), the reliability of the specific facets and the importance ratings could not be determined. However, when Marsh, Barnes and Hocevar (in press) measured specific components of self-concept with both multi-item scales and single-item responses, they concluded that the single-item responses were not an adequate substitute. More generally, Rushton, Brainerd, and Pressley (1983) argued that single-item measurements are less stable, less reliable, less valid, and less generalizable than multi-item scales.

Scaling the Self-concept and Importance Ratings. Hodge and McCarthy concluded that group values were more important than individual values, but they actually found that neither group nor individual ratings of importance were useful. None of the weighted averages based on any of the importance ratings was substantially more correlated with self-esteem than was the unweighted average of specific components. While the weighted average based on individual importance ratings did poorest of all, this may be an artifact of using unscaled scores. Since self-concepts are multiplied by importance ratings, a self-concept scale with a higher mean response will be more affected by differences in important ratings than if all self-concept scores were standardized (i.e., mean = 0, SD = 1). Also, the unscaled importance ratings may overemphasize the mean importance rating by each subject across all facets, and underemphasize the differences in rating made by a given subject. For example, consider subjects who attach high importance to facets where their self-perceptions are good and low importance to facets where their self-perceptions are poor. The interactive hypothesis predicts that these subjects should have a very good Global Esteem, but the weighted average of the specific facets would be substantially lower than if these subjects rated all facets -- even the ones where their self-perceptions were poorer -- to be important. It is proposed that a set of ipsative weights based on the importance ratings (i.e., a set of weights that sum to a constant for each subject) will produce weighted averages that are more positively correlated to esteem than the unscaled importance ratings used by Hodge and McCarthy.

The Role of Specific and General Facets. Implicit in the Hodge and McCarthy study is the suggestion that the role of specific facets is to define a general facet, so that if the specific facets do not correlate

highly with the general facet they are seen to be less useful. An alternative perspective consistent with the new emphasis on the multidimensionality of self-concept is that if the general facet is not substantially correlated to each of the specific facets, then the general facet should be seen as less useful.

The Present Investigation.

The purpose of the present investigation is to relate global self-esteem to specific facets of self-concept and the importance of each facet. The study is based on data from all published studies by the author that have used the SDQ III. Hence, the multiple dimensions of self-concept are based on a well articulate model, the Shavelson model, and the psychometric properties (dimensionality, reliability, validity) of the instrument are very good. Self-concept facets are represented by both multi-item scales and single-item responses, allowing a comparison between the two. Diverse subgroups who have completed the SDQ III are expected to differ widely in the importance they attach to different facets (e.g., Physical Ability for a group of athletes, and Spiritual Values in students from a senior Catholic girls school), providing a good basis of comparing the influence of individual, subgroup, and group values. Analyses are performed with both scaled and unscaled importance ratings, and with standardized and unstandardized self-concept scores, thus providing a test of the proposed artifact in the Hodge and McCarthy study. Though there are many differences between the present investigation and the Hodge/McCarthy study, the Global Esteem scale from the SDQ III was initially derived from the Rosenberg scale that was used in the earlier study and many of the specific facets considered in the two studies are similar.

METHODS

The Self Description Questionnaire (SDQ) III.

The present investigation is based on responses to the SDQ III. The SDQ III, the rationale for its construction, its relation to the Shavelson model and the other SDQ instruments, the wording of the items, exploratory and confirmatory factor analyses identifying the facets that it is designed to measure and its hierarchical structure, the internal consistency and stability of its scales, and its relation to academic achievement, to self-concept inferred by significant others, and to other constructs is summarized elsewhere (Marsh, 1984; Marsh, Barnes & Hocevar, in press; Marsh & Jackson, 1984; Marsh & O'Neill, 1984; Marsh, Richards & Barnes, in press). Each of the 13 SDQ III scales is represented by 10 or 12 items, approximately half of which are negatively worded, and subjects respond on an eight-point response scale where categories vary from "1-Definitely

"False" to "8-Definitely true." The 13 scales and a summary item are:

1. Physical Ability -- I am good at sports and physical activities.

2. Physical Appearance -- I am physically attractive/good looking.

3. Opposite Sex Relations -- I have good interactions/relationships with members of the opposite sex.

4. Same Sex Relations -- I have good interactions/relationships with members of the same sex.

5. Relations With Parents -- I have good interactions/relationships with my parents.

6. Spiritual Values -- I am a religious/spiritual person.

7. Honesty -- I am an honest, reliable, trustworthy person.

8. Emotional Stability -- I am an emotionally stable person.

9. Verbal -- I have good verbal skills/reasoning ability.

10. Math -- I have good mathematical skills/reasoning ability.

11. General Academic -- I am a good student in most school subjects.

12. Problem Solving -- I am good at problem solving/creative thinking.

13. General-Self (called Global Esteem for purposes of this study) -- I have self-respect, self-confidence, self-acceptance, positive self-feelings and a good self-concept.

Though not formally part of the SDQ III, these studies have also asked subjects to respond to 12 summary description items (the first 12 presented presented above) designed to reflect 12 of the 13 scales -- all but the Global Esteem scale. For each of these 12 items subjects indicate the item's accuracy (i.e., how accurate is this statement as a description of you) and its importance (i.e., how important is this characteristic in determining how you feel about yourself). Responses to these items are made on a "1-very inaccurate/very unimportant" to "9-very accurate/very important" scale. Psychometric properties of the accuracy ratings and their relation to the multi-item scale scores that they are designed to reflect were examined by Marsh, Barnes and Hocevar (in press), and are further considered as part of the present investigation. There has been no previous investigation of properties of the importance ratings or how they relate to the different areas of self-concept for the SDQ III.

For purposes of the present investigation the following variables derived from the SDQ III responses are considered:

1) Raw Self-concept scale scores -- the raw responses to the 10 or 12 items designed to measure each scale, after reverse scoring the negatively worded items, were summed and divided by the number of items in the scale.

2) Standardized Self-concept scale scores -- the same as 1 except that

scale scores were standardized to have mean = 0, SD = 1 across the total group.

3) Raw Summary Ratings -- the raw responses to the accuracy ratings for the 12 summary description items that measure different facets of self-concept.

4) Standardized Summary ratings -- the same as 3 except that scale scores were standardized to have mean = 0, SD = 1 across the total group.

5) Raw Importance Ratings -- the raw responses to the importance ratings for the 12 summary description items.

6) Proportionalized Importance ratings -- for each subject each importance rating was divided by the sum of the 12 importance ratings such that the transformed scores varied between 0 and 1 (actually .01 and .45 since each facet had an importance between 0 and 9). Hence, for each individual subject the set of 12 transformed scores summed to 1.0 (i.e., the scores were ipsative) and had a mean of 1/12.

7) Global Esteem -- the mean of responses to the 12 Global Esteem items.

The Samples.

For purposes of the present investigation data are considered from five distinguishable groups of subjects that are described elsewhere and are briefly summarized below:

1) Group 1 consists of the sample of 151 Australian university students (mean age = 21.9, 79% female) described by Marsh, Barnes and Hocevar (in press; also see Marsh & O'Niell, 1983, study 2). As part of that study, subjects also asked the person in their life who knew them the best to complete the SDQ III. These significant others completed the survey as if they were the person who had given it to them (i.e., they were to predict what the subject had said). For purposes of the present investigation, the primary focus is on the self ratings but the self-other correlations are also considered.

2) Group 2 consists of the 361 Outward Bound participants (mean age = 21.3, 24% females) described by Marsh, Richards and Barnes (in press). As part of that study, participants completed the SDQ III one month before, the first day of, and the last day of a 26-day residential program. For purposes of the present investigation, the primary focus is on data from time 1, but test-retest correlations over the time 1-time 2 interval are also considered.

3) Group 3 consists of the 296 year-11 girls (mean age = 16.7) from two private catholic girls schools described by Marsh and O'Niell (1983).

4) Group 4 consists of 46 high-school girls (mean age = 16.6) who were selected as the "nonathlete" control for subjects summarized in group 5 (see Marsh and Jackson, 1984; Jackson, 1984). Unlike subjects in groups 1 - 3,

subjects in groups 4 and 5 completed accuracy and importance ratings for the 12 summary items, but completed only four of the self-concept scales (Physical Ability, Physical Appearance, Emotional Stability, and Global Self-esteem). For purposes of the present investigation, only the summary rating items and their relation to General Esteem are considered for these two groups. Nevertheless, results from the original study showed that subjects in the nonathlete group, compared to the athletes, were substantially lower in self-concept of Physical Ability and the importance they placed on this facet, but did not differ in terms of other areas of self-concept or the importance placed on these areas. Hence, these groups are particularly relevant to the purposes of the present investigation.

5) Group 5 consisted of 76 women athletes (mean age = 20.1) consisting of 46 high school women athletes and 30 finalists in the 1984 Australian women's powerlifting championships (Marsh and Jackson, 1984; Jackson, 1984), and these subjects completed the same materials as those in group 4. Though the high school athletes and powerlifters differed slightly in of some areas of self-concept in a way that appeared to be age-related, the differences were small and they did not differ in terms of ratings in the Physical Ability facet or any of the importance ratings. For this reason, and also because of the small sample sizes, both groups of women athletes were combined to form group 5 for purposes of the present investigation.

Statistical Analysis

All statistical analyses were performed with the commercially available Statistical Package for the Social Sciences (SPSS; Nie, et al., 1975; Hull & Nie, 1981). Relations between the self-concept scales and Global Self-esteem were performed across subjects in groups 1 - 3 and separately for each group, while relations between the summary ratings and Global Self-esteem were conducted across all subjects in groups 1 - 5 and separately for each group. Except for the self-concept scales not completed by groups 4 and 5, there was very little missing data for any of the subjects and mean responses were substituted for the few missing values that did occur. The actual analyses are described in more detail in the results section.

It was anticipated that the diverse subgroups who have completed the SDQ III would vary substantially in the importance they placed on different areas. In preliminary analyses, multivariate and univariate analyses confirmed these expectations, and these data are summarized in Table 1. Because of the large sample size virtually every effect is statistically significant, but the etas from one-way ANOVAs provide an estimate of the size of each effect. A detailed examination of these group differences is not the focus of this study, but of particular relevance is the

demonstration that the groups do vary substantially and logically in the importance they place on some facets -- particularly the Physical Ability and Spiritual Values facets. These preliminary findings also provide support for the validity of the importance ratings when averaged across a large number of respondents.

Results and Discussion.

Psychometric Properties of Responses to the Summary and Importance Ratings.

Single-item responses tend to have limited reliability, validity, and generalizability that may affect conclusions based on them. Internal consistency coefficients are the most easily obtained and frequently used estimates of reliability, but they cannot be computed on the basis of a single score. Consequently studies that use single-item responses often disregard reliability and its implications. However, in the present investigation test-retest correlations were be used to estimate the reliability of single-item responses and their validity was examined by relating them to external criteria.

Insert Table 2 About Here

Subjects in group 2 completed the self-concept scales, the summary ratings, and the importance ratings twice during a one month interval, and the test-retest correlations appear in Table 2. Coefficients for the multi-item scales (median r = .87) are substantial and similar in magnitude to the internal consistency estimates that were reported in the original study. The coefficients for the single-item summary ratings designed to parallel the multi-item scales are substantially lower (median r = .70), but still moderate. However, test-retest coefficients for the importance ratings (median r = .57) are so low that their usefulness may be dubious.

Significant others were selected by subjects in group 1, and these significant others inferred responses (i.e., completed the SDQ III as if they were the subject) to the self-concept scales, the summary ratings, and the importance ratings. Correlations between the self-reports and the responses by the significant others were used to test the validity of responses to the SDQ III (see Marsh, Barnes & Hocevar, in press, for more detail). The self-other correlations (Table 2) are substantial for the multi-item self-concept scales (median r = .57), lower for the summary self-concept ratings (median r = .32), and lowest for the importance ratings (median r = .19). These findings provide support for the validity of the self-concept scales and summary ratings, but little support for the validity of importance ratings.

Correlations between each of the 12 multi-item self-concept scales and the corresponding single-item summary rating were also computed (Table 2). Not surprisingly, these correlations are substantial, particularly given the

apparent unreliability of the summary ratings. These findings suggest that the summary ratings may provide a reasonable estimate of the multi-item scales that they are designed to parallel, even though they are apparently less reliable and valid (see Marsh, Barnes & Hocevar, in press, for further discussion).

The apparent lack of reliability and validity of the importance ratings is discouraging, particularly since they are the main focus of the present investigation. However, there is a wide variation in the estimates for the different facets. In particular, the test-retest and self-other correlations are substantially higher for the Spiritual facet and to a lesser extent the Physical Ability facet than for the other facets (self-other agreement on the self-concept scales and summary ratings are also highest for these two facets). Hence, while interpretations based upon the importance ratings must be made cautiously, there is evidence that ratings for some of the facets may have acceptable psychometric properties. It is also important to recognize that the poor psychometric properties of the importance ratings refer only to individual responses, and the the mean importance rating across a large group of subjects may be substantial even when the reliability of individual responses is only modest.

Relations Among Specific Facets, Importance Ratings and Global Esteem.

Self-concept/Importance Correlations. Hodge and McCarthey described, but did not test Rosenberg's selectivity hypothesis (1965; 1982) that posits that individuals attach higher importance to the specific facets where their self-perceptions are the most positive. Correlations between the importance ratings and the self-concept measures (Table 2), both the multi-item scales (median r = .43) and the summary ratings (median r = .35), support this selectivity hypothesis. However, neither the present study nor Rosenberg's research, tested the direction of the causal relationship that is ~~implicit~~ in the selectivity hypothesis (i.e., that higher self-perceptions "cause" higher importance ratings), and it seems plausible that subjects may strive to improve their skills and self-perceptions in areas that they view to be important. It is interesting to note that again the coefficients for the Spiritual and, to a lesser extent, the Physical Ability facets are higher than the other facets. Consistent with earlier observations, these may also those facets where subjects have the most freedom in determining the relative importance that is attached to them and where the importance ratings are most variable. In summary, these findings demonstrate that specific facets of self-concept are moderately correlated with the importance attached to each area, and that for some facets this relationship is very strong.

Insert Table 3 About Here

Unweighted Averages of Specific Self-concepts. Three unweighted averages were used to summarize responses to the 12 specific facets of self-concept -- the average of the raw responses, the average of standardized responses, and the first principal component[1] derived from the set of responses. Each of these unweighted averages was correlated with Global Esteem for the total group and each subgroup (Table 3). For the multi-item scales all the unweighted averages correlate about .7 with Global Esteem, while correlations for the summary ratings are close to .6. Though there are minor differences among the three unweighted averages, and among the different subgroups, the results are remarkably consistent.

These correlations based on unweighted averages provide a lower limit for testing the usefulness of the various weighted averages. To the extent that the weighted averages perform no better than the unweighted averages, than the rule of parsimony dictates that the unweighted averages are preferable. Hodge and McCarthy (1984) reported a correlation of only .41 between their unweighted average and Global Esteem, a value substantially lower than even the .6 value found for the summary ratings in the present investigation. While many possible explanations exist for the different results, the most likely are the reliability of the Global Esteem measure, the age of the subjects (which is related to the reliability of the Rosenberg scale in the Hodge & McCarthy study), and the number of specific components that are considered. Responses to the Global Esteem scale from the SDQ III are more reliable than responses to the Rosenberg scale in the Hodge and McCarthy study. Hodge and McCarthy found that the specific components were more highly correlated for the Global Esteem for their oldest subjects, perhaps due to differences in the reliability in their Global Esteem measure that were age-related, and subjects in the present study were older than subjects in their study. Hodge and McCarthy considered only 9 specific components, compared to the 12 used here, and so their unweighted average was probably less reliable.

Specific Facets Weighted By Group and Subgroup Importance Ratings. For these analyses each self-concept score was weighted by the mean importance weighting either for the entire group, or for the subgroup to which the subject belonged (the means appear in Table 1). While different procedures for scaling the specific self-concept and importance ratings were considered, these make virtually no difference in the results and so are not discussed further. For the multi-item self-concept scales, correlations between the weighted averages and Global Esteem are close to .7, while those for the summary items are close to .6. In each instance, averages weighted by total group importance ratings are virtually identical to those

weighted by subgroup importance ratings, and do not differ from those based on unweighted averages. The correlations reported here are higher than the Hodge and McCarthy study, probably for the same reasons as indicated above, but the similarity of correlations based on weighted and unweighted responses is consistent with their findings. These findings provide no support for the usefulness of the group or subgroup importance ratings in the weighting of specific components of self-concept to predict Global Esteem.

Empirically Weighted Averages. A series of multiple regressions were conducted to predict Global Esteem from the set of 12 self-concept scales and from the set of 12 self-concept summary ratings. For the total group the multiple Rs (Table 3) are .78 and .67 for the self-concept scales and summary ratings respectively. Similar analyses were performed for each subgroup, and these multiple Rs are somewhat higher. However, multiple Rs are somewhat biased and the extent of this bias depends on the number of variables used in the prediction equation and the number of cases. When the multiple Rs were corrected for this bias (see Nie, et al, 1975), differences between the multiple Rs for each subgroup and for the total group are much smaller (Table 3)[2].

The multiple regressions for the total group represent an absolute upper limit for correlations based on averages weighted by the total group ratings of importance. Similarly, the multiple regressions for each subgroup represent the absolute upper limit for weighed averages based on the subgroup ratings of importance. While these multiple Rs are clearly higher than those obtained with the unweighted averages, the size of the differences are modest. Since the difference between correlations derived from the empirically determined optimal weights and from the unweighted average is less than .1, it is not possible for any set of a priori weights such as those based on the importance ratings to do substantially better than the unweighted averages.

Specific facets Weighted By Individual Importance Ratings. For these analyses each self-concept facet for each subject was weighted by the importance rating assigned to that facet by the subject. These analyses are particularly important in that: a) this is the type of weighting typically implied by researchers who advocate the use of importance ratings; b) this is the type of weighting that Hodge and McCarthy found to perform poorer than the unweighted average and which was suggested might be an artifact of scaling problems in earlier discussions; and c) the empirically determined multiple R does not constitute an absolute upper limit for the correlation between this weighted average and Global Esteem[3].

Unlike the previous analyses, the way that the self-concept scores and

the importance ratings were scaled makes a substantial difference (Table 3). When raw self-concept scores were weighted by raw importance ratings, as in the Hodge and McCarthy study, the weighted averages were substantially less correlated with Global Esteem than any of the other weighted or unweighted averages. These findings are consistent for both self-concept scales and summary ratings, are consistent across subgroups, and are consistent with Hodge and McCarthy's findings. However, when either self-concept scores are standardized or the importance ratings are proportionalized, the correlations are substantially larger. These correlations are also marginally higher than other nonempiricially determined weightings.

In order to further explore the apparent artifact of using unscaled scores, each subject was assigned a set of 12 random numbers that varied between 0 and 1,[4] and these were used to weight the specific self-concepts. Except for the use of these random numbers in place of the importance ratings, the same analyses were performed. Again the correlations based on the raw (random) weights and the raw self-concept scores result in substantially lower correlations than when either the self-concept scores are standardized or the weights are proportionalized. This demonstrates that the findings are not due to some serendipitous relationship between the importance ratings and self-concept scores that is idiosyncratic to the present investigation. It is also gratifying to note that the randomly weighted averages are somewhat less correlated with Global Esteem than previously considered averages, though even these differences are surprisingly small.

Two general conclusions result from this set of analyses based on individual weightings. First, as posited, the Hodge-McCarthy conclusion that these weighted averages perform more poorly than any other weighted or unweighted average is apparently an artifact of using unscaled scores[5]. Second, when the scores are scaled, the weighted averages based on individual importance ratings perform marginally better than do the unweighted averages, the averages weighted by group importance ratings, and the averages weighted by subgroup importance ratings. Are these marginal differences large enough to be theoretically important? -- probably not in that the improvement never exceeds .05 for scales or summary ratings, for any subgroup, or for any way of computing the weighted and unweighted averages. Hence, these analyses again fail to provide much support for the usefulness of importance ratings in the weighting of specific components of self-concept to predict Global Esteem.

The Rosenberg Interactive Hypothesis.

The Rosenberg interactive hypothesis described earlier posits that the

effect of a specific self-concept on Global Esteem will vary with the value that an individual places on the specific facet. In the language of analysis of variance (ANOVA), the level of self-concept in a particular facet and the facet's importance will interact in determining Global Esteem, though not all interactions would be inconsistent with the hypothesis. Two different approaches were taken in testing this model as described below:

1) A Classical ANOVA Model. Scores for each specific self-concept facet and for each importance rating for the total group were divided into four categories (i.e., low, medium-low, medium-high, high). Using Global Esteem as the dependent variable, a 4 (level of specific self-concept) x 4 (level of importance) ANOVA was conducted separately for each facet using the classical approach for testing statistical significance (see Nie, et al., 1975). The F-ratios for each effect, the two main effects and the interaction effect, and the multiple Rs based on the variance explained by all three effects are summarized in Table 4.

2) A Multiple Regression Model: The ANOVA model described above provides a general test of each of the effects, but the particular form of significant effects may not be consistent with the Rosenberg hypothesis. Consequently, a more specific test was devised. In this approach, only three parameters were estimated instead of the 15 in the classical ANOVA model (i.e., the 15 df in the 4 x 4 ANOVA): the linear effect of the specific facet of self-concept, the linear effect of its importance, and the effect of the product of the two (i.e., the linear-by-linear effect in the interaction term). For purposes of this analysis raw self-concept scores and proportionalized importance ratings were used rather than the categorized scores used to test the classical ANOVA model. Standardized beta weights and pearson correlations representing each effect, as well as the multiple R based on the three effects are summarized in Table 4.

The large number of statistical tests and differences between the two approaches make the results of these analyses difficult to summarize. It is useful to compare the multiple Rs from the ANOVA and multiple regression models to each other and to the simple correlation between the specific self-concept and General Esteem. The multiple R based on the classical ANOVA model might be larger than the multiple R for the regression model since more degrees of freedom are used in the estimation process, but it could be lower if substantial variance is lost in the categorization of scores. However, if the multiple Rs do not substantially exceed the simple correlation between the specific self-concept facet and General Esteem (i.e., if the additional terms do not contribute to the variance explained), then there is no support for the Rosenberg hypothesis. Inspection of the

results shows that multiple Rs based on the multiple regression model are slightly higher than those for the ANOVA model, providing support for the more specific test of the Rosenberg interaction model. However, few of the multiple Rs derived from this model are substantially larger than the coresponding pearson correlation. Consistent with previous analyses, the largest differences between the multiple Rs and the pearson rs occur for the Spiritual and Physical Ability facets.

Results described thus far provide tests for of Rosenberg interactive hypothesis for each specific facet of self-concept, but the multiple regression model is easily extended to test the model across all 12 facets. A hierarchical regression was performed in which the set of 12 self-concept scores was entered on the first step, the set of importance ratings was entered on the second step (see footnote c in Table 4), and the set of 12 interaction terms was entered on the third step. Statistical tests (Table 4) indicated that the interaction effects do contribute significantly to the prediction of Global Esteem beyond the effects of specific self-concepts, but the added contribution is small. While the interpretation of interaction effects due to a specific facet may be problematic, the exclusion of the interaction effect for the Spiritual facet produced a larger decrease in the multiple R than any other interaction term, and this finding was consistent across analyses of the self-concept scales and summary ratings.

In summary, support for the Rosenberg model is weak but not nonexistent. While the contribution of the self-concept/importance interaction is small, it is statistically significant. It also appears that support of the model is stronger for the Spiritual and perhaps the Physical Ability facets, and this is intuitively logical and consistent with previous observations.

Summary and Implications

The Interactive Hypothesis.

A variety of theoretical hypotheses, as well as common sense, posit that the effect of a specific facet of self-concept on general self-concept will depend on the facet's importance. William James first proposed the hypothesis 100 years ago; it has been restated frequently; it seems intuitive plausible; and it should be easy to test. Nevertheless, rigorous tests of the hypothesis -- indeed, even clearly articulated accounts of how it should be tested -- are surprisingly rare. Hodge and McCarthy, in one of the most recent attempts, found little or no support for the hypothesis, but an examination of their study suggested methodological problems. The present investigation verified that the methodological problems did exist, but even when these were rectified there was little support for the hypothesis. The interactive hypothesis has too much intuitive appeal to be

completely rejected, and so further examination of the issues is needed.

One potential problem in the present investigation was the use of single-item responses to assess the importance of each facet. The importance ratings were psychometrically weak, and asking subjects to rate the importance of each item in the muli-item scales would provide better scores. However, none of the weightings considered in the present study, not even the random weighting, really made much difference. Hence, it seems unlikely that obtaining better importance ratings would substantially alter the weighted averages or improve support for the hypothesis.

Ironically, another possible problem is that the SDQ III is designed too well in that it only includes facets that are at least reasonably important to most people and it represents most of the facets considered to be important by the respondents. None of the facets was given a mean importance rating of less than 5, the middle response category, and all but one had means between 6 and 8 on a nine-point response scale. Even if a particular subject attached small importance to one or two of these facets, this would have only a small impact on the weighted average since most of the other facets would be judged to be important. In contrast to the broadly defined, generally important facets that appear on the SDQ III, the two facets originally suggested by William James to illustrate this proposal were skills as a psychologist, his profession, and skills at Greek language, an area in which he had "no pretensions." Consistent with James' original formulation, support for the proposal would probably be much stronger if the "facets" were comprised of narrowly defined characteristics that most respondents judged to be unimportant but that a few found to be very important. Of the SDQ III facets, the Spiritual and Physical Ability facets seem closest to this description, and they also had the most variable importance ratings. It is also these facets that provided the best support for the interactive proposal. While a new instrument such as suggested here might provide better support for the interactive hypothesis, it would not make a very good self-concept instrument.

Other Issues.

Several other issues, though not the primary focus of the present investigation, were also addressed. These include the use of single-item scales, the status of the general self construct, and the relation between self-perceptions and the value placed on specific facets of self-concept.

Single-item scales designed to represent the SDQ III facets were found to have poorer psychometric properties than the multi-item scales, and this is consistent with findings in other areas or research and common-sense. Nevertheless, there was modest support for the psychometric properties of

the summary ratings. Thus, while it is better to use the 136 SDQ III items than the 12 summary items, it may be better to use than 12 summary items than to not to consider multiple dimensions of self-concept at all. If external constraints preclude the use of the entire SDQ III, then the use of the most relevant subscales and the set of 12 summary rating items may be an expedient compromise.

According to the selectivity hypothesis, subjects selectively assign the greatest importance to those areas of self-concept in which their self-perceptions are most positive. While this hypothesis is often linked to the interaction hypothesis, the two are separate. In the present investigation, there was clear support for the selectivity hypothesis even though support for the interaction hypothesis was very weak. However, further research is needed to test the assumption of causality that is implicit in the selectivity hypothesis.

Implicit in the Hodge and McCarthy study, and in the theoretical underpinnings that stimulated the present study, is the assumption that general self-concept is more important than specific facets; the role of specific self-concepts has been to provide a basis for defining general self-concept. However, this may be a historical weakness in self-concept research. Elsewhere (Marsh & Shavelson, 1984) I have argued that self-concept cannot be adequately understood if its multidimensionality is ignored. If the role of self-concept research is to better understand the complexity of the self, to predict diverse behaviors, and to relate self-concept to other constructs, then measures of multiple facets are more useful than a general facet. Though specific facets such as those measured by the SDQ III are hierarchically ordered, the hierarchy is so weak that the general self that appears at the apex accounts for only a small portion of the variance in the specific facets (Marsh, 1984). This situation appears to be analogous to the one in intelligence/ability testing where many researchers now argue against the sole reliance on IQ. While the general constructs in both areas of research may be relevant for the very young, specific facets become more important as individuals grow older and the specific facets become more differentiated. Ironically, since the hierarchy of specific facets in self-concept research appears to be weaker than in intelligence/ability research, the reliance on a general construct is less justifiable in self-concept research even though it appears to be more prevalent.

FOOTNOTES

1 -- Principal component analyses, with 1.0's as communality estimates, were performed by the commercially available SPSS program (Nie, et al., 1975) on the set of 12 self-concept scales and the set of 12 self-concept summary ratings. While these are not really "unweighted" averages, they are considered as such in this discussion because the weights do not depend on either the importance ratings or the Global Esteem score. For the 12 scales the factor score coefficients (see Table 1 for the ordering the the 12 facets) are: .15, .18, .18, .17, .14, .19, .00, .12, .19, .11, .19 and .20. For the 12 summary ratings the factor score coefficients (in the order) are: .14, .15, .17, .15, .13, .17, .03, .16, .19, .16, .19 and .19.

2 -- The standardized beta weights for the 12 self-concept scales (in the same order as in Table 1) are: -.08, .25, .18, .16, .08, .36, .10, .08, -.02 - .06, .08 and .17. Those for the 12 summary ratings (in the same order) are: .01, .23, .12, .14, .11, .33, .01, .02, .09, - .04, .00 and .08.

3 -- The multiple R based on the total group represents the optimal weighting of the 12 self-concept facets when the weight assigned to each facet is the same for all subjects, but not when the weight assigned to each facet is allowed to vary from subject to subject. Hence, it is possible for the weighted averages based on individual weights to outperform the multiple regression prediction though this did not happen.

4 -- Random numbers from a uniform distribution varying between 0 and 1 were generated with the SPSS procedure (Nie, et al., 1975) and were used to weight the self-concept facets. This analysis was replicated five times with different sets of random numbers and the results were quite consistent over the different replications. The results that appear in Table 3 are actually the mean correlations resulting from the five replications.

5 -- The difference between the proportionalized and raw weights is only a concern for averages weighted by the individual importance ratings. When the total group means were used as weights, the proportionalized and raw weights were perfectly correlated, and when subgroup means were used the two sets of weights were almost perfectly correlated. The findings support the contention that the paradoxically low correlation between individually weighted self-concept facets and Global esteem is an artifact of using unscaled scores, but the precise nature of this artifact is not clear to me. I anticipated that the use of standardized self-concept scales would produce a small improvement, that the use of ipsative importance ratings would produce a larger improvement, and that the use of both would do the best, but this is not what I found. The fact that the problem disappears when either proportionalized importance ratings or standardized self-concept scores are used to determine the weighted averages suggests that the problem

is not inherent in either set of scores. Since the same pattern was observed when randomly assigned weights were used, explanations that involve response biases in the way subjects make importance ratings are not viable. Since the finding generalizes across self-concept scales and summary items here, and apparrantly the summary items used by Hodge and McCarthy suggests that it is not a function the wording of the self-concept measures. (I suspect that I can come up with a reasonable explanation, perhaps one suggested by the journal reviewers, before this paper is actually published, and it will provide a more satisfactory ending to this footnote.)

REFERENCES

Boersma, F.J. & Chapman, J.W. (1979). Student's Perception of Ability Scale Manual. University of Alberta. Edmonton, Canada.

Burns, R. B. (1979). The self-concept: Theory, measurement, development and behaviour. London: Longman.

Coopersmith, S. (1967). The antecedents of self-esteem. San Francisco: Freeman.

Dusek, J. B., & Flaherty, J. F. (1981). The development of self-concept during the adolescent years. Monographs of the Society for Research in Child Development, 46(4, Serial No. 191).

Fleming, J. S. & Courtney, B. E. (1984). The dimensionality of self-esteem: II: Hierarchical facet model for revised measurement scales. Journal of Personality and Social Psychology, 46, 404-421.

Harter, S. (1982). The Perceived Competence Scale for Children. Child Development, 53,87-97.

Harter, S. (1983). Developmental perspectives on the self-system. In P. H. Mussen (Ed.), Handbook of Child Psychology, (Volume IV, 4th edition, pp. 275-385). New York: Wiley.

Hodge, D. R. & McCarthy, J. D. (1984). Influence of individual and group identity salience in the global self-esteem of youth. Journal of Personality and Social Psychology, 47, 403-414.

Hull, C. H., & Nie, N. H. (1981). SPSS Update 7-9. New York: McGraw-Hill.

Jackson, S. (1984). Athletic or antisocial: The female sport experience. Psychological determinants and consequences of female sports involvement. An unpublished thesis completed as partial fulfillment of requirements for the degree of B. Ed. (Hons) in Physical Education, University of Sydney.

James, W. (1890/1963). The principles of psychology. New York: Holt, Rinehart & Winston.

Marsh, H. W. (1984). The hierarchical structure of self-concept: An application of confirmatory hierarchical factor analysis. (in review). Department of Education, University of Sydney, Australia.

Marsh, H. W., Barnes, J., Cairns, L., & Tidman, M. (1984). The Self Description Questionnaire (SDQ): Age effects in the structure and level of self-concept for preadolescent children. Journal of Educational Psychology, 76, 940-956.

Marsh, H. W., Barnes, J. & Hocevar, D. (in press). Self-other agreement on multidimensional self-concept ratings: Factor analysis & multitrait-multimethod analysis. Journal of Personality and Social Psychology.

Marsh, H. W., & Jackson, S. (1984). Multidimensional self-concepts, masculinity and femininity as a function of women's involvement in athletics. (In review).

Marsh, H. W., & O'Niell, R. (1984). Self Description Questionnaire III (SDQ III): The construct validity of multidimensional self-concept ratings by late-adolescents. Journal of Educational Measurement, 21, 153-174.

Marsh, H. W. & Parker, J. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? Journal of Personality and Social Psychology, 47, 213-231.

Marsh, H. W., Richards, G., & Barnes, J. (in press). Multidimensional self-concepts: The effect of participation in an Outward program. Journal of Personality and Social Psychology.

Marsh, H. W. and Shavelson, R. J. (1984). Self-concept: Its multifaceted, hierarchical structure. Department of Education, University of Sydney, Australia. (in review)

Nie, N. H., Hull, C. H., Jenkins, J. G. Steinbrenner, K. & Bent, D. H. (1975). Statistical Package for the Social Sciences. New York: McGraw-Hill.

Rosenberg, M. (1982). Psychological selectivity in self-esteem formation. In M. Rosenberg & H. B. Kaplan (Eds.), Social psychology of the self-concept (pp. 535-546). Arlington Heights, Ill: Harlan Davidson.

Rosenberg, M. (1965). Society and the adolescent self-image. Princeton: Princeton University press.

Rosenberg, M. (1979). Conceiving the self. New York: Basic Books.

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. Psychological Bulletin, 94, 18-38.

Shavelson, R. J. & Bolus, R. (1982). Self-concept: The interplay of theory and methods. Journal of Educational Psychology, 74, 3-17.

Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Validation of construct interpretations. Review of Educational Research, 46, 407-441.

Shavelson, R. J., & Marsh, H. W. (In Press). On the structure of self-concept. In R. Schwarzer (Ed.), Anxiety and cognitions. New Jersey: Erlbaum.

Soares, L. M., & Soares, A. T. (1982). Convergence and discrimination in academic self-concepts. Paper presented at the 20th Congress of the International Association of Applied Psychology, Edinburgh, Scotland, July, 1982.

Watkins, D. (1978). The development and evaluation of self-esteem measuring instruments. Journal of Personality Assessment, 42, 171-182.

Welles, L. E., & Marwell, G. (1976). Self-esteem: Its conceptualization and measurement. Beverly Hills, Calif.: Sage Publications.

Wylie, R. C. (1974). The self-concept (Rev. ed., Vol. 1) Lincoln: University of Nebraska Press.

Wylie, R. C. (1979). The self-concept (Vol. 2) Lincoln: University of Nebraska Press.

Table 1

Mean Ratings (and SDs) For Self-concept Scales (Scale), Summary Self-Concept Ratings (Sum) and Importance Ratings (Imp) and Eta's representing Differences Between Groups on These Scores.

		Facets[a] 1	2	3	4	5	6	7	8	9	10	11	12	13
Total Group														
Scale	Mn	5.72	4.78	4.32	4.80	5.82	5.18	4.87	6.11	5.47	4.89	5.59	5.11	5.71
	SD	1.42	1.23	1.11	0.84	1.20	1.25	1.66	0.72	1.06	1.56	1.10	0.96	1.16
Sum	Mn	5.67	5.05	5.75	6.56	6.65	6.41	4.75	7.06	6.27	5.63	6.08	5.85	---
	SD	1.91	1.59	1.85	1.47	1.85	1.73	2.42	1.27	1.44	1.87	1.47	1.53	---
Impt	Mn	5.96	5.95	7.01	7.13	7.66	7.34	5.34	7.97	7.01	6.08	6.72	6.63	---
	SD	2.03	1.82	1.53	1.50	1.52	1.48	2.52	1.25	1.45	1.75	1.56	1.62	---
Group 1														
Scale	Mn	5.42	5.18	4.70	4.80	5.62	5.15	4.90	6.17	5.75	4.54	5.83	5.06	6.00
SD	SD	1.53	0.99	0.99	0.85	1.34	1.26	1.90	0.74	1.05	1.70	0.93	0.95	1.05
Sum	Mn	5.33	5.75	6.41	6.59	6.33	6.24	4.95	7.13	6.54	5.20	6.20	5.94	---
	SD	2.01	1.35	1.68	1.49	1.88	1.85	2.60	1.27	1.40	2.02	1.53	1.36	---
Impt	Mn	5.47	6.40	7.29	7.27	7.64	7.33	5.49	7.93	7.10	5.49	6.93	6.63	---
	SD	2.12	1.62	1.54	1.43	1.59	1.56	2.64	1.24	1.36	1.76	1.32	1.53	---
Group 2														
Scale	Mn	6.13	5.25	4.28	4.67	5.91	5.40	4.41	6.17	5.52	5.10	5.68	5.36	5.73
SD	SD	1.07	1.03	1.08	0.92	1.14	1.20	1.75	0.73	1.05	1.56	1.19	0.96	1.19
Sum	Mn	6.16	5.40	5.72	6.32	6.60	6.53	4.33	7.17	6.32	5.89	6.19	5.95	---
	SD	1.61	1.29	1.77	1.50	1.77	1.70	2.45	1.12	1.37	1.87	1.55	1.46	---
Impt	Mn	6.61	5.90	7.24	6.99	7.52	7.34	4.83	8.04	7.14	6.24	6.47	6.87	---
	SD	1.65	1.70	1.32	1.43	1.48	1.43	2.58	1.22	1.31	1.58	1.57	1.50	---
Group 3														
Scale	Mn	5.28	4.04	4.17	4.96	5.81	4.92	5.41	6.02	5.25	4.81	5.37	4.82	5.53
SD	SD	1.53	1.16	1.18	0.70	1.19	1.21	1.18	0.71	1.04	1.45	1.04	0.89	1.10
Sum	Mn	5.17	4.29	5.34	6.84	6.82	6.28	5.49	6.84	5.92	5.40	5.84	5.57	---
	SD	1.98	1.71	1.91	1.29	1.78	1.64	1.94	1.17	1.35	1.74	1.24	1.53	---
Impt	Mn	5.13	5.64	6.60	7.32	7.87	7.21	6.28	7.86	6.64	6.01	6.83	6.24	---
	SD	2.07	1.97	1.62	1.42	1.46	1.53	1.94	1.27	1.54	1.81	1.55	1.69	---
Group 4														
Scale	Mn	--	--	--	--	--	--	--	--	--	--	--	--	5.27
SD	SD	--	--	--	--	--	--	--	--	--	--	--	--	1.31
Sum	Mn	4.04	4.54	5.37	6.33	6.85	6.26	4.09	6.78	6.37	5.50	6.00	5.80	---
	SD	1.51	1.59	1.81	1.71	2.11	1.77	2.66	1.81	1.60	1.86	1.56	1.66	---
Impt	Mn	5.46	6.28	7.24	7.15	7.78	7.91	4.78	8.13	7.47	6.54	7.26	6.85	---
	SD	1.94	1.71	1.23	1.63	1.79	1.05	2.91	1.24	1.38	2.05	1.69	1.67	---
Group 5														
Scale	Mn	--	--	--	--	--	--	--	--	--	--	--	--	5.96
SD	SD	--	--	--	--	--	--	--	--	--	--	--	--	1.20
Sum	Mn	6.97	5.29	6.46	6.67	6.71	6.80	3.92	7.34	6.80	6.15	6.37	6.37	---
	SD	1.44	1.59	1.75	1.60	2.16	1.91	2.56	1.69	1.71	1.80	1.57	1.83	---
Impt	Mn	7.45	6.26	6.90	6.82	7.46	7.50	4.20	8.13	7.43	6.55	6.76	6.87	---
	SD	1.46	1.94	1.89	1.98	1.91	1.48	2.55	1.34	1.66	1.75	1.75	1.79	---
Eta's[b]														
Scale		.29	.47	.17	.16	.09	.18	.27	.10	.17	.14	.16	.25	.14
Sum		.36	.36	.23	.16	.09	.10	.24	.14	.19	.17	.12	.15	---
Impt		.39	.15	.20	.12	.11	.10	.28	.08	.19	.18	.14	.17	---

a -- 1=Physical Ability; 2=Physical Appearance; 3=Opposite Sex Relations; 4=Same Sex Relations; 5=Parent Relations; 6=Emotional; 7=Spiritual Values; 8=Honesty; 9=Verbal; 10=Math; 11=General Academic; 12=Problem Solving; 13=Global Esteem.

b -- Oneway ANOVAs of differences between groups are summarized in terms of the eta values. For the scale scores the eta's are based on 808 sets of responses from groups 1-3 (eta > .08 is statistically significant at $p < .05$). For summary and importance ratings the eta's are based on the 930 sets of responses from groups 1-5 (eta > .09 is statistically significant at $p < .05$).

Table 2

Correlations Between Self-concept Scales (Scale), Summary Ratings (Sum), and Importance Ratings (Imp) for 12 Areas of Self-concept.

Area	Test-Retest rs[a]			Self-Other rs[b]			rs Between[c]		
	Scale	Sum	Impt	Scale	Sum	Impt	Scale & Sum	Scale & Impt	Sum & Impt
Physical	.88	.77	.61	.77	.65	.45	.81	.59	.62
Appearance	.85	.72	.65	.47	.48	.18	.73	.23	.06
Opposite Sex Relations	.89	.76	.55	.51	.31	.33	.80	.37	.29
Same Sex Relations	.85	.70	.61	.45	.32	.20	.65	.48	.27
Parents	.88	.70	.66	.76	.55	.19	.77	.54	.43
Emotional	.85	.66	.50	.62	.28	.19	.64	.28	.11
Spiritual	.94	.84	.86	.82	.77	.70	.87	.85	.86
Honesty	.75	.54	.50	.37	.15	.19	.49	.44	.36
Verbal	.88	.66	.46	.62	.33	.09	.62	.35	.13
Math	.91	.80	.57	.75	.63	.23	.83	.43	.46
General Academic	.87	.69	.57	.36	.22	.26	.68	.36	.26
Problem Solving	.86	.68	.49	.46	.15	.17	.68	.47	.36
General[d]	.88	---	---	.45	---	---	---	---	---
Median r	.87	.70	.57	.57	.32	.19	.71	.43	.35

a -- Test-retest correlations (interval = 1 month, n=361) from Marsh, Richards and Barnes (in press).

b -- Self-other correlations (n=151) are based on self-report responses and self-concepts inferred by a significant other from Marsh, Barnes & Hocevar (in press; also see Marsh & O'Neill, 1984, Study 2).

c -- correlations based on the Total Sample (n=930 for Sum and Impt ratings and 808 for Scales)

d -- The General Self was only assessed by scales.

Table 3

Correlations Between Weighted Averages of the 12 Self-concept Scores (Scales and Summary Ratings) and General Self for the Total Group and Each Subgroup

	rs relating the weighted average of scale scores to Esteem for Group:				rs relating the weighted average of summary ratings to Esteem for Group:					
	Total	1	2	3	Total	1	2	3	4	5
[a] Unweighted Averages										
Raw Score	.673	.658	.699	.646	.586	.551	.631	.518	.579	.559
Standard Score	.688	.678	.708	.666	.591	.560	.635	.525	.587	.552
Principal Component	.702	.707	.715	.695	.588	.565	.635	.522	.583	.541
[b] Weighted By Total Group Importance Ratings										
Raw Score/Raw Impt	.685	.672	.709	.659	.595	.564	.641	.532	.585	.559
Raw Score/Prop Impt	.685	.672	.709	.659	.595	.564	.641	.532	.585	.559
Stand Score/Raw Impt	.693	.683	.712	.671	.595	.568	.641	.535	.591	.540
Stand Score/Prop Impt	.693	.683	.712	.671	.595	.568	.641	.535	.591	.540
[c] Weighted By Subgroup Importance Ratings										
Raw Score/Raw Impt	.680	.676	.710	.656	.593	.568	.643	.530	.588	.554
Raw Score/Prop Impt	.685	.677	.711	.650	.596	.568	.643	.530	.589	.554
Stand Score/Raw Impt	.693	.686	.713	.669	.597	.572	.642	.534	.593	.544
Stand Score/Prop Impt	.693	.686	.713	.669	.597	.572	.643	.534	.594	.544
[d] Empirically Weighted Importance Ratings										
Uncorrected rs	.777	.790	.800	.778	.673	.677	.721	.660	.836	.762
Corrected rs	.773	.768	.792	.767	.669	.640	.709	.641	.768	.708
[e] Weighted By Individual Importance Ratings										
Raw Score/Raw Impt	.514	.480	.523	.503	.478	.439	.514	.432	.454	.455
Raw Score/Prop Impt	.701	.700	.722	.670	.615	.595	.664	.544	.632	.549
Stand Score/Raw Impt	.703	.693	.718	.683	.606	.584	.653	.540	.609	.555
Stand Score/Prop Impt	.706	.698	.722	.680	.611	.588	.658	.543	.633	.538
[f] Weighted By Individual Randomly Assigned Weights										
Raw Score/Raw Impt	.379	.293	.392	.360	.378	.358	.410	.337	.347	.308
Raw Score/Prop Impt	.632	.620	.657	.602	.557	.542	.600	.483	.551	.524
Stand Score/Raw Impt	.652	.652	.672	.621	.560	.559	.589	.486	.531	.513
Stand Score/Prop Impt	.657	.654	.677	.629	.567	.559	.606	.492	.568	.524

a -- Unweighted averages of the 12 self-concept scores (scales and summary ratings) were computed taking the average of the raw scores, standardized scores (Stand score), and the first principal component. Each was then correlated with the General Self score.

b -- Weighted scores are defined as in b except that the important weightings were obtained the total group importance ratings rather than from individual importance ratings.

c -- Weighted scores are defined as in b and c except that the important weightings were obtained the subgroup ratings rather than from individual or total group importance ratings.

d -- Optimal empirical weights were obtained from a series of multiple regressions performed for the total group and each subgroup. Corrected correlations are adjusted for the number of variables included in the multiple regression, and this adjustment varies with the sample size (see Nie, et al., 1975). The multiple regression based on the 12 scales and the 12 summary ratings was .796 (.789 when corrected). When General Self scores were predicted from importance ratings from each subgroup separately, the General Self score correlated .796 (.781) with the weighted average of scale scores and .721 (.781) with the weighted average of summary ratings.

e -- Raw scores and standard (Stand) scores were weighted by importance ratings assigned by each individual. Separate averages were derived from the raw importance ratings (Raw Impt) and the proportional importance (Prop Impt) ratings.

f -- Weighted scores are defined as in b except that random variables (that varied between 0 and 1 on a uniform distribution) were used instead of the importance ratings.

Table 4

The Relation Between Each Self-concept Score (scales and summary ratings), Each Importance Rating, and the Self-concept/Importance Interaction on General Self

Area			Effect of: Self-concept	Importance	Interaction	Multiple R[a]
Physical	Scale	ANOV[b]	37.34**	6.29**	2.01*	.33**
		MR	-.05 (.27)	-.85** (.03)	1.0** (.18)	.37**
	Sum	ANOV[c]	23.4**	3.13*	2.59**	.27**
		MR	.00 (.25)	-.51** (.03)	.64** (.19)	.32**
Appearance	Scale	ANOV	91.96**	8.04**	1.59	.49**
		MR	.41** (.51)	-.23 (-.09)	.16 (.28)	.53**
	Sum	ANOV	23.4**	3.13*	2.59**	.46**
		MR	.28** (.41)	-.32** (-.10)	.23 (.24)	.45**
Opposite Sex Relations	Scale	ANOV	80.41**	8.81**	1.82	.50**
		MR	.08 (.50)	-.45** (.01)	.64** (.39)	.52**
	Sum	ANOV	55.52**	8.11**	1.22	.41**
		MR	.25** (.41)	-.16* (.01)	.07 (.32)	.43**
Same Sex Relations	Scale	ANOV	59.12**	1.38	1.40	.43**
		MR	-.21 (.46)	-.87** (-.01)	1.15** (.30)	.50**
	Sum	ANOV	43.65**	5.53**	3.34**	.36**
		MR	.15 (.35)	-.39** (-.01)	.41* (.23)	.39**
Parent	Scale	ANOV	24.34**	0.99	1.46	.30**
		MR	.18 (.34)	-.23 (.04)	.28 (.25)	.35**
	Sum	ANOV	27.03**	2.43	1.54	.30**
		MR	.42** (.30)	.07 (.04)	-.09 (.23)	.32**
Emotional	Scale	ANOV	127.32**	1.23	1.67	.55**
		MR	.56** (.60)	-.02 (.08)	.06 (.50)	.60**
	Sum	ANOV	88.06**	2.26	0.77	.48**
		MR	.49** (.51)	-.04 (.08)	.03 (.43)	.51**
Spiritual	Scale	ANOV	6.41*	0.47	0.84	.17*
		MR	-.07 (.06)	-.50** (-.01)	.60** (.05)	.17**
	Sum	ANOV	2.44	1.16	2.34*	.12**
		MR	-.03 (.08)	-.41** (-.01)	.47 (.08)	.19**
Honesty	Scale	ANOV	12.22**	1.69	2.28*	.23**
		MR	.16 (.23)	-.13 (.03)	.15 (.15)	.23**
	Sum	ANOV	16.29**	1.01	1.84	.23**
		MR	.29* (.28)	-.04 (.04)	.01 (.22)	.29**
Verbal	Scale	ANOV	27.74**	1.00	0.79	.32**
		MR	.30** (.34)	-.06 (.01)	.07 (.26)	.34**
	Sum	ANOV	41.02**	0.81	1.21	.35**
		MR	.30* (.34)	-.12 (.01)	.08 (.25)	.34**
Math	Scale	ANOV	5.04**	3.91**	1.51	.16**
		MR	.16 (.12)	-.16 (-.06)	.04 (.06)	.18**
	Sum	ANOV	7.48**	2.47	1.98*	.16**
		MR	.03 (.14)	-.26** (-.06)	.25 (.08)	.19**
General Academic	Scale	ANOV	29.83**	1.98	0.88	.32**
		MR	.48** (.32)	.03 (.00)	-.21 (.19)	.34**
	Sum	ANOV	23.69**	1.82	0.89	.27**
		MR	.19 (.25)	-.12 (.00)	.07 (.17)	.26**
Problem Solving	Scale	ANOV	56.27**	6.09**	1.76	.42**
		MR	.48** (.41)	-.14 (-.02)	-.04 (.23)	.43**
	Sum	ANOV	38.54**	3.99**	3.78**	.33**
		MR	.13 (.31)	-.35** (-.01)	.37 (.21)	.35**

Table 4 continued

Mult R for Scale[d]	.785**	.791**	.836**
Change in Mult R Squared	.616**	.009	.073**
Multiple R for Sum[d]	.673**	.687**	.704**
Change in Mult R	.453**	.019**	.024**

a -- The effects of each area of self-concept, the corresponding importance rating, and their interaction on General Self was analyzed in series of 4 (levels of self-concept) by 4 (level of importance) ANOVAs. The reported values are the F-ratios for each effect and the multiple R is the square root of the SSexplained/SStotal ratio in each analysis. The df for error terms in analyses of the scales is approximately 778, and that for analyses of the summary ratings is approximately 909.

b -- A series of multiple regressions (MR) was conducted where each self-concept, the corresponding importance rating (proportion), and the product of these two variables was used to predict Global Esteem. The standardized beta weight for each effect is presented for this analysis, along with the simple correlation (in parentheses) between each variable and General Self, and the multiple R for all three variables.

c -- These Multiple Rs resulted from a hierarchical multiple regression in which the 12 self-concept scores were entered first, followed by the 11 importance ratings (since the proportionalized importance ratings were used, any one is a linear combination of the other 11 and one had to arbitrarily be excluded), and then the 12 products obtained by multiplying each self-concept by its corresponding importance rating. Significance tests were conducted for the Multiple R at each step, and for the change in Multiple R squared for each step.